METEOROLOGICAL OFFICE
125455
25 NOV 1977
LIBRARY

MET O 11 TECHNICAL NOTE No. 99

Errors in Rainfall Forecasts by the 10-Level Model

1973 - 1977

---

by

P G Wickham

Meteorological Office (Met O 11)
London Road
Bracknell
Berkshire
United Kingdom

November 1977

NB. This paper has not been published. Permission to quote from it should be obtained from the Assistant Director of the above Meteorological Office branch.

Errors in Rainfall Forecasts by the 10-Level Model - 1973 - 1977 - P G Wickham

1. The data.

   Period
   Region
   Forecast rainfall
   Actual rainfall

2. The forecast errors.

   Distribution of individual errors
   Magnitude of mean monthly error
   Mean monthly error
   Overall ratio of forecast/actual rain
   Consistency of the rainfall predictions on successive forecast runs.

3. Evidence for changes in forecast skill during the period.

Rainfall Forecasts by the 10-Level Model - 1973-1977

Since August 1973, the 12-hourly rainfall accumulations predicted by the 10-level model (rectangle version) have been compared with the corresponding actual accumulations, for a region covering part of the land area of W Europe. This Note records the performance of the model in terms of monthly mean error statistics over sub-regions approximately 4 x 4 grid-lengths in size (1.6×10^5 km^2).

1. The data

Period: 47 months, from August 1973 to June 1977 inclusive

Region: The whole region, sub-divided into 10 sub-regions, is shown in Fig 1. The sub-regions, numbered 3-12, were chosen to be roughly similar in size and to be areas within which there is broadly similar climate and geography. They are listed in Table I.

TABLE I.

Sub-region		Number of forecast grid points	Normal number of actual rainfall reports
Number	Description		
3	Denmark	17	13
4	Scotland	17	19
5	N.England	11	13
6	S.England	13	19
7	Ireland	14	14
8	Holland	16	29
9	Germany	19	18
10	Belgium	12	14
11	W.France	23	11
12	E.France	22	9
	AVERAGE	16.4	15.9

The grid-points of the 10-level model are approximately 100 km apart, so the size of an average sub-region, containing 16 grid-points is 1.6×10^5 km^2. The whole region, containing 164 grid-points covers about 1.3×10^6 km^2.

Forecast rainfall: The 10-level model (rectangle) is run operationally every 12 hours, starting from midnight and midday data. The forecast is carried forward to 36 hours (T+36), so that each 12-hour period of real time is covered by part of the forecast on three successive runs, as displayed in Fig 2.

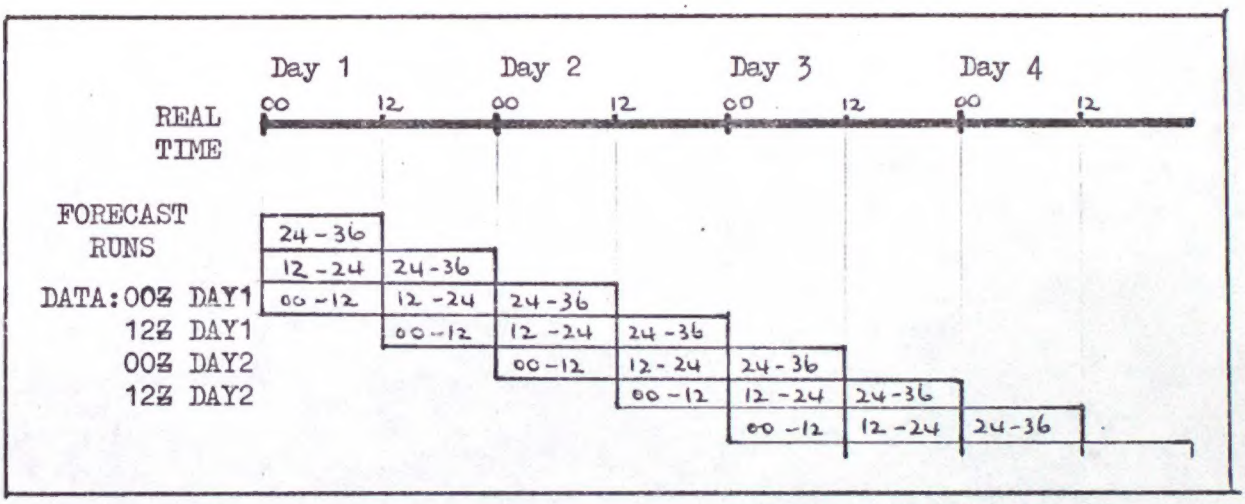

Fig:2.

The rainfall (R) accumulated at each grid-point of the 10-level model is the sum of the accumulations resulting from two processes and are normally referred to as "dynamic rain" (D) and "convective rain" (C). During the period being considered there have been various changes to the formulation of the physical processes in the model which have altered the balance between D and C, but there has been little change to the overall total R (= D+C), which is the quantity with which this Note is concerned.

No rainfall occurs at the initial time-step of a forecast run, and the forecast rain is very deficient throughout the first 6-9 hours of each run. This is the length of time required for the model's fields of moisture and vertical motion to become fully developed from the

smooth, initial fields from which the forecast run is started. No attempt has been made in this Note to use the accumulations from the first 12 hours of a forecast run.

Each 12-hour accumulation of actual rain can therefore be associated with two forecasts:-

the (24-36) hour forecast from one run
and the (12-24) hour forecast from the following run.

In this Note the mean of these two forecasts has been used as the Forecast Data, and these may conveniently be referred to as "24-hour forecasts". The forecast rainfall for an area is the mean of the forecasts at each of the individual grid-points within the area.

There are some systematic changes in the model's overall behaviour during a forecast run. The deficiency in rainfall during the first 12-hour period has already been mentioned, and beyond this time the (24-36) forecasts show an increase of 6% over the (12-24) forecasts in the total amount of forecast rain. The magnitude of the mean monthly errors (irrespective of their sign) also increases with time, rising from 0.77 to 0.88 mm/12 hours between the (12-24) and (24-36) periods - an increase of 14%. At the same time the mean error (taking account of its sign) becomes rather closer to zero, indicating some increase in the number of occasions of over-forecasting of rain in the later stages of the forecasts, although the overall tendency to underforecast rain is still very pronounced.

Actual rainfall: The actual rainfall data used in this work have been derived from the land station synoptic reports stored, daily, in the Met O 22 International Archive Data-set on the COSMOS computer

system. The normal number of 12-hourly rainfall reports available from each sub-region are shown above in Table I. The spatial density of this network is not high, being about one observation to a 100 km square. Bearing in mind the typical gradients which occur in the distribution of rainfall, an observational network should be much denser than this for one isolated observation to give a reliable indication of the average rainfall in its surrounding region. However the four year sample which is available should be large enough to allow acceptable comparisons to be made between actual and forecast rainfall, on the space and time scales used in this Note.

In general, the actual rainfall for a sub-region has been taken to be the mean of the individual accumulations at all the available reporting stations in the sub-region, though slight departures from this principle have sometimes occurred in practice. For example, the data acquisition program performed some limited quality control of the observational data, to the extent of not including individual 12-hour accumulations of more than 30 mm in the calculated areal mean values. Such reports were then printed out for inspection, and were only used if supporting evidence of high areal rainfall was available from another source (such as the DWR in the case of British Stations) or from consideration of the synoptic situation. Also - for those occasions when either the computer or the archiving system broke down, areal mean values for the 4 British sub-regions (4-7) were subsequently prepared by hand, using DWR data. This has occurred on average with a frequency of 3 runs/month.

Some broad generalisations about the character of the rainfall over the region are relevant to the interpretation of forecast performance.

(a) In summer the rainfall over most of the region is predominantly convective in character. Over the continental sub-regions this is certainly so. And even over those sub-regions which experience some Atlantic cyclonic activity, many of the fronts are strongly influenced by surface heating and the development of convective instability.

(b) In winter the rainfall is largely frontal, with a much smaller proportion of convective activity. The western sub-regions particularly, come under the influence of oceanic frontal systems, which are unmodified by surface heating and convective activity is confined to exposed coastal regions in polar air-masses. Fronts penetrate right into the continent also, where orography is the main modifying influence, rather than convection.

(c) During the particular 4 year period to which this Note refers there have been some unusual features in the temporal distribution of rainfall. Much of the period has been quite dry over most of the area, and the well remembered 'drought' over much of the UK during 1975/6 is a significant part of the data sample. Another feature, clearly seen in Fig 4(a), has been the occurrence of a series of wet autumns which, coupled with the dry seasons in between, has produced a cyclic character to the annual precipitation curve.

TABLE II Frequency of occurrence (%) of 12-hour areal accumulations of specified amounts

12-HR RAINFALL		0	Tr	1-2	3-5	6-9	10 or more	(mm)
DESCRIPTION		NIL	TRACE	SLIGHT	MODERATE	HEAVY	VERY HEAVY	
MONTHS	WET	21	18	36	17	7	1	100
WHICH	AVERAGE	29	25	34	9	3	.	100
ARE	DRY	63	17	16	3	1	.	100

where WET/DRY values are from the 3 wettest/3 driest months in the period and AVERAGE values are from the 3 months with rainfall closest to the period mean.

The frequency of occurrence of areal rainfalls having specified values is markedly skew. Table II, which defines the descriptive terms (slight, moderate etc) used in this Note, also shows that even in the wettest months nearly 40% of the forecasts are either dry or produce negligible amounts of rain. In the dry months this figure rises to 80%. Consequently a good yardstick against which to compare the numerical forecasts is a no-skill forecast which always predicts dry conditions.

2. The forecast errors.

(a) The distribution of individual errors.(f-a)

Figure 3 shows the overall distribution of the occurrence of forecast errors, summed over the whole region and over the whole period - it is based on over 24000 individual forecasts.

It shows that over 80% of the forecasts have been less than 2 mm in error. By any standard these must be judged to be good forecasts, though a large proportion of them are occasions when no rain fell and none was forecast. However it does show that a very high proportion of the numerical forecasts can be described as giving useful guidance, as shown by the impressive clustering of the distribution about the

central 'zero' value. Unfortunately the rather long, flat tails to the distribution indicate that there are 5-10% of the forecasts which go astray, and maybe badly astray.

The occasions which give rise to large errors are necessarily such that either a lot of rain was forecast which did not occur or a lot fell that was not forecast. Either way those occasions are likely to have been significantly important events, for which forecasters would have been particularly anxious to have good advice from the numerical model. It is on these comparatively rare, but very important, occasions of significant rainfall that the practical value of the numerical forecasts has to be assessed just as much as on the 80% of occasions when both the human forecaster and the numerical model can achieve good forecasts in non-developmental, and essentially trivial forecasting situations.

TABLE III The Percentage Frequency of Forecast and Actual Rain

		Forecast Rain						
		0	Tr	Slt	Mod	Hvy	V.H.	
ACTUAL RAIN	0	27.7	5.5	3.1	0.1	.	.	36.4
	Tr	7.9	4.9	3.9	0.4	.		17.1
	Slt	6.5	9.2	14.3	2.3	0.3	.	32.6
	Mod	0.6	1.4	5.5	2.3	0.5	.	10.3
	Hvy	0.1	0.2	1.3	1.0	0.5	0.1	3.0
	V.H	.	.	0.2	0.2	0.2	.	0.6
		42.7	21.2	28.2	6.3	1.4	0.2	100.0

WHOLE REGION
WHOLE PERIOD

Table III shows how the degree of success of rainfall forecasts is distributed when the rainfalls are grouped into 6 categories (see Table II for definitions). The general tendency to under-forecast

the wet occasions and over-forecast the dry ones can be seen. Also, the success rate at predicting the occasions when moderate or heavy rain actually occurred is quite low.

When the individual errors are summed over p periods and s sub-regions (where $p \gtrsim 50$ or more) they show some regular, non-random characteristics which are conveniently summarised by the seasonal curves displayed in Fig.4. The rather more detailed curves of the corresponding features on a monthly time-scale are in Fig 5. These display the same regular characteristics as the seasonal curves.

(b) <u>The magnitude of the mean monthly error</u> $\left(\frac{1}{ps}\sum|f_{ps}-a_{ps}|\right)$

The absolute magnitude of the mean forecast error, over the whole region, during periods of a month or more, shows a very close correlation with the actual rainfall. The two curves (Fig.4(a) and (b)) are remarkably similar, the forecast errors being large in wet seasons, and small in dry seasons. Using monthly data the correlation coefficient between the two quantities is 0.97. Thus the dependance on actual rainfall accounts for over 90% of the variability of the magnitude of the errors from month to month.

The regression line between E*, the mean magnitude of the error (in mm) of a forecast 12-hour accumulation and r, the average actual 12-hour rainfall (approx 1/60 of the monthly total) is

$$E^* = 0.13 + 0.7r$$

The normal value of r is about 1 mm (lying between the extremes 0 and 2 mm), so the magnitude of the forecast errors averaged over a period of a month is between $\frac{1}{2}$ and $\frac{3}{4}$ of the actual rainfall total.

(c) The mean monthly error $\left[\frac{1}{ps}\sum(f_{ps}-a_{ps})\right]$

The mean error has a negative bias, indicating that the forecast rain is on average less than the actual rain. Of course there are occasions when the numerical model predicts more rain than actually occurs but such occasions are consistently outweighed by errors in the opposite sense.

There is some association between the value of the mean monthly error and the amount of the monthly rainfall (Figs 4(c) and (a)) but the correlation ($r = -0.56$) is not close. The largest mean errors (strongly underforecast) have occurred in the wet autumn months with some regularity, but the association at other times of the year is less exact.

The mean errors have come closest to zero during the winter seasons, when Atlantic cyclonic systems are the main rain-producing agents. Also, winter is the season when the proportion of successfully forecast occurrences of moderate or heavy rainfalls reaches its peak:

TABLE IV

Autumn (96/192)	= 0.50
Winter (123/151)	= 0.81
Spring (49/93)	= 0.53
Summer (26/98)	= 0.27
Annual (294/534)	= 0.55

Proportion of actual rainfalls $\geq$ moderate intensity which are correctly forecast.

There are doubtless many ways in which individual occurrences of over-forecasting can arise so as to affect these particular statistics (e.g. through spurious developments; or slight discrepancies between forecast and actual tracks of lows, particularly as between a sea or

a land track). One incorrect mode of development which results in over-forecasting of rain in some areas and which has been observed on a number of occasions is associated with the development of an increasingly meridional flow in the E Atlantic, on the western border of the region considered in this Note. In these situations, when the W'ly zonal flow over the atlantic starts to buckle then the eastward movement of cyclonic systems across the British Isles is slowed down. The main Atlantic cyclonic vortex becomes slow-moving to the west of Ireland and the flow over Britain becomes SW'ly or S'ly. The numerical model has on a number of occasions been very slow at predicting this increasing meridionality and as a consequence has predicted a too markedly eastward motion of significant frontal rain belts from the Atlantic into British and continental areas which in fact remain largely dry.

(d) <u>The overall ratio of forecast/actual rainfall in the region</u> $\left(\Sigma f_{ps}/\Sigma a_{ps}\right)$

This ratio takes no account of whether the forecast rain was predicted to fall in the right place at the appropriate time, it is simply a measure of the total monthly rainfall that accrued over the region (both actual and forecast).

The ratio shows a well-marked seasonal variability and little or no dependance on the actual rainfall amount. The immediate inference of this is that it reflects the way in which the numerical model handles the different physical processes responsible for producing rain during the year.

In summer, the actual rainfall is predominantly convective in character and the numerical predictions are serious under-estimates (only 45% of the rain is forecast), regardless of whether it predicts the location correctly or not. This must be considered to be a consequence (partly

at least) of the imperfect parameterisation of the sub-grid-scale convective rainfall processes used in the formulation of the numerical model. In passing it may be noted that the introduction of the 'deep convective' parameterisation late in 1975 appears to have done nothing to improve the value of this ratio during the subsequent summer, 1976 (which was, of course, one of the outstandingly dry periods of recent times). What the new scheme did do was to increase, very markedly, the amount of "convective rain" predicted by the model, but this was offset by a corresponding decrease in the amount of "dynamic rain" predicted - so the moisture budget of the model in summer remains seriously deficient.

In winter, the convective element in the actual rainfall diminishes in importance, even in continental areas. The rain comes mainly from oceanic frontal systems. In these situations the model predicts more nearly the correct quantity of overall rainfall (about 85% of the rain is forecast). Thus it seems that the grid-length (100 km) and the dynamical equations of the model are such that they can handle the developments in synoptic-scale rain systems reasonably well. This was indeed the type of rainfall which the 10-level model was originally designed to investigate.

(e) The consistency of the rainfall predictions on successive forecast runs

In day to day synoptic forecasting a high level of consistency in the advice given by numerical forecasts on successive runs is most desirable. It is consistency in performance, as much as the absolute accuracy of the results, which determines how quickly and how completely forecasters are able to build up their confidence in the usefulness of the numerical products.

Experience has shown that there is only a rather weak tendency for the rectangle rainfall forecasts either to give improved predictions as the lead time decreases, or to show any other consistent properties that can be reliably documented and allowed for by forecasters. There seem to be unpredictable effects which swamp any regularity in its performance. It is possible that these effects stem largely from the scattered distribution and limited representativeness of the particular data used in successive forecast runs, especially the oceanic humidity data.

To illustrate the overall consistency of successive forecasts, Table V(a) shows that on 62% of all occasions the same category of rainfall is predicted for a given 12-hour period in a particular sub-region by successive forecasts. (This includes 'no rainfall' as one of the categories). On 31% of occasions the forecasts differ by one category and only on 7% of occasions is there a substantial difference of two categories or more. A slight skewness is apparent in the Table, due to the somewhat larger rainfall amounts predicted at (24-36) hours than earlier in the forecast run.

Table V(b) is a compression of Table V(a), displaying the consistency of simple Dry/Wet discriminations. The figures are encouraging, particularly because a bias towards dry forecasts might have been expected from the model's tendency to underforecast rainfall. In fact the proportions of both 'dry' and 'wet' forecasts at (24-36) hours which are confirmed by a later, (12-24) hours forecast is exactly the same (83%).

However, when attention is focussed on the many fewer, but very important, occasions of significant rainfall - in Table V(c) -

TABLE V Comparison of (12-24)forecasts and (24-36)forecasts for the same period

(a) Forecast rainfall in six categories

		(24-36) Forecasts						
		0	Tr	Slt	Mod	Hvy	V.H.	
(12-24) Forecasts	0	342	74	26	2	.	.	444
	Tr	54	99	64	5	1	.	223
	Slt	16	53	150	30	4	1	254
	Mod	1	4	24	22	7	1	59
	Hvy	.	1	3	6	4	1	15
	V.H		.	1	1	1	1	4
		414	231	268	67	17	3	1000

(b) Discriminating 'Dry' (No Rain) and 'Wet' (Some Rain) Occasions

		(24-36) Forecasts		
		Dry	Wet	
(12-24) Forecasts	Dry	342	102	444
	Wet	72	484	556
		414	586	1000

(c) Discriminating 'Very Wet' (Moderate Rain at least) and 'Not Very Wet' (Dry or Slight Rain)

		(24-36) Forecasts		
		NVW	VW	
(12-24) Forecasts	NVW	879	43	922
	VW	34	44	78
		913	87	1000

the consistency with which these events are forecast is relatively much less. Only half the 'very wet' forecasts at (24-36) hours are confirmed by a similar forecast 12 hours later. And since the relative frequencies of:

actual occurrence/(12-24) forecast/(24-36) forecast = 13/8/9

for these occasions of significant rain, the overall usefulness of the numerical predictions for these events is rather low.

3. Evidence for changes in forecast skill during the 4-year period

Because the size of the monthly forecast error depends so strongly on the wetness of the month, the simplest way of displaying variations in the level of forecast performance is to express the magnitude of the error as a proportion of the monthly rainfall. Figure 6 illustrates this statistic, which has been called the Error Ratio. Improvements in the forecast skill are indicated by smaller values of the Error Ratio.

It can be seen from Figure 6 that there is no obvious evidence of any consistent improvement in the forecasts over the period as a whole. Though from time to time it has been suggested that some features of the Error Ratio curve may be attributable to specific causes, the only conclusion that can really be reached from the evidence of Figure 6 is that there is no discernible change in the average Error Ratio throughout the four year period.

This is a very disappointing result, in view of the efforts which have been made in a number of ways to improve the standard of the rainfall forecasts during this time. Figure 6 also shows the dates of various changes to the operational forecast system which could have had some effect on the quality of the forecasts. These changes include:

(a) changes to the oceanic data distribution -
especially the withdrawal of the US weather ships by the end of 1974 and the start of the revised N Atlantic Weather Ship agreement in mid-1975.

(b) changes in the formulation of the forecast model -
especially the introduction of a deep-convection scheme in June 1975 and the revised lateral boundary changes brought into effect in April 1976.

(c) changes in the operational running of the numerical forecasts -
especially the practice of running the rectangle before the octagon, introduced in August 1974, and the necessary use of boundary changes from the previous octagon run also the use of octagon background fields for the rectangle analyses, introduced in Jan 1975.

The last mentioned change was probably the one which in practice was observed to have the most clearly beneficial effect, as it stabilised the quality of the forecast products and increased their reliability. Prior to Jan 1975 the rectangle forecasts appeared to suffer from a lack of regular operational monitoring. The effects of uncorrected errors in the analyses were perpetuated from one run to the next through poor background fields. This gave rise to an erratic standard of performance, which may have contributed to the oscillations shown by the Error Ratio curve during 1973-74.

The introduction of background fields from the (well-monitored) octagon at the start of 1975 was accompanied also by an expansion in the numbers of monitoring staff. It was tempting to look with some

satisfaction at the resulting Error Ratio curve during the 15 months from Jan 1975 to March 1976. It displayed remarkable steadiness and had a definite improving trend during this period.

However the character of the curve changed sharply during the summer of 1976. This was an extremely dry period in the UK and much of western Europe so that the impact of the deteriorating forecasts was quite small. Part of the reason for the poor performance during this period was subsequently thought to have been explained by the discovery that incorrect humidity background fields had been used. This accidentally came about after changes were made to the organisation of the operational computer programs. The close association between the period when incorrect fields were in use and the months when the Error Ratio rose to high values is striking. Soon after the operational procedures were corrected the Error Ratios in the autumn of 1976 reverted for a while to their previous level.

The subsequent increase in the Error Ratio during 1977 has been disappointing and no specific reasons for it can be adduced. Investigations have been made which have eliminated as far as possible some of the sources of error encountered in the past - such as inadvertant operational changes to the organisation or programming errors introduced with new formulations of the model The absence of any specific explanation in this case probably demonstrates that the forecasting skill does not depend in a clearcut way on any single causative factor. It may also indicate that the Error Ratio, as applied here to mean monthly values and test areas which are inadequately defined by actual observations, is too insensitive a statistic to give any reliable information.

Whatever the reason may be, the results of this investigation do not demonstrate that there has been any overall change of any kind in the quality of the numerical precipitation forecasts over the past four years. This conclusion is similar to that reached by Fawcett (1977) about the numerical forecasting capability in the United States.

REFERENCE:

FAWCETT E B 1977 "Current capabilities in prediction at the National Weather Services N.M.C" Bull:Amer:Met:Soc: Vol 58 No.2. pp 143-149

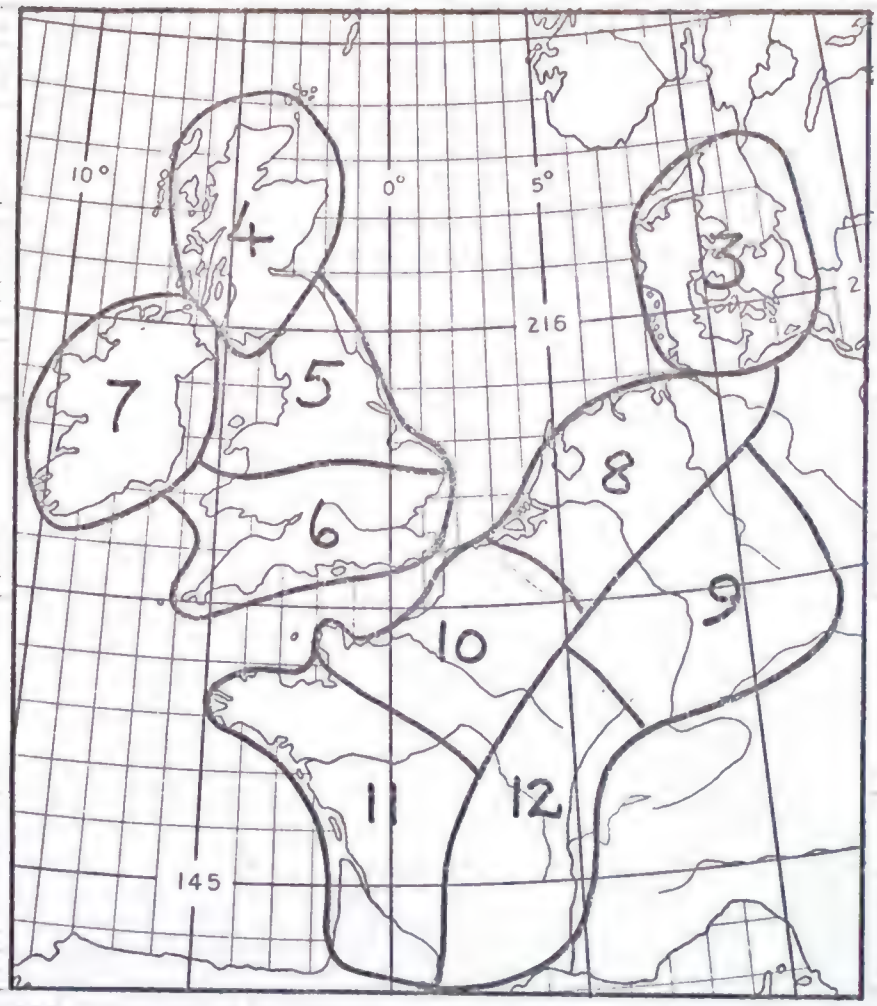

Figure 1.

The sub-regions, with their identification numbers.(see Table I).

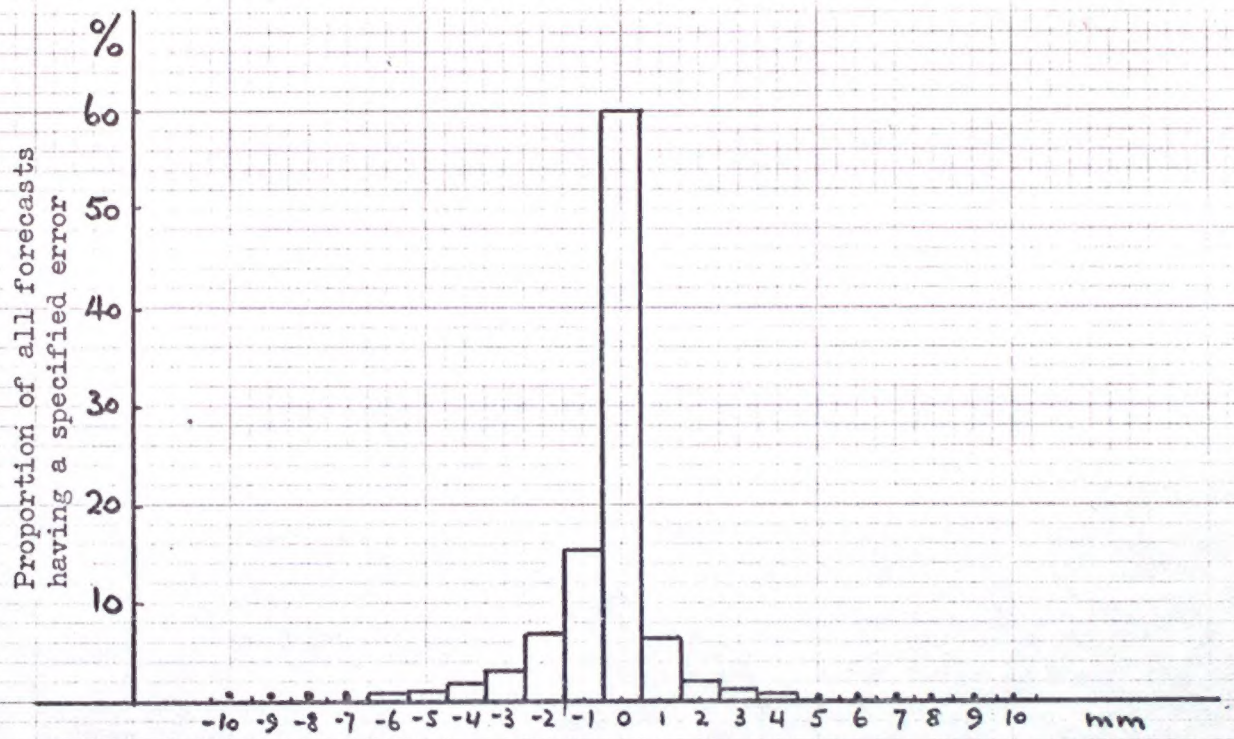

Figure 3.

The overall distribution of forecast errors, for all sub-regions throughout the whole period.
The data used here are (12 - 24) hour forecasts.

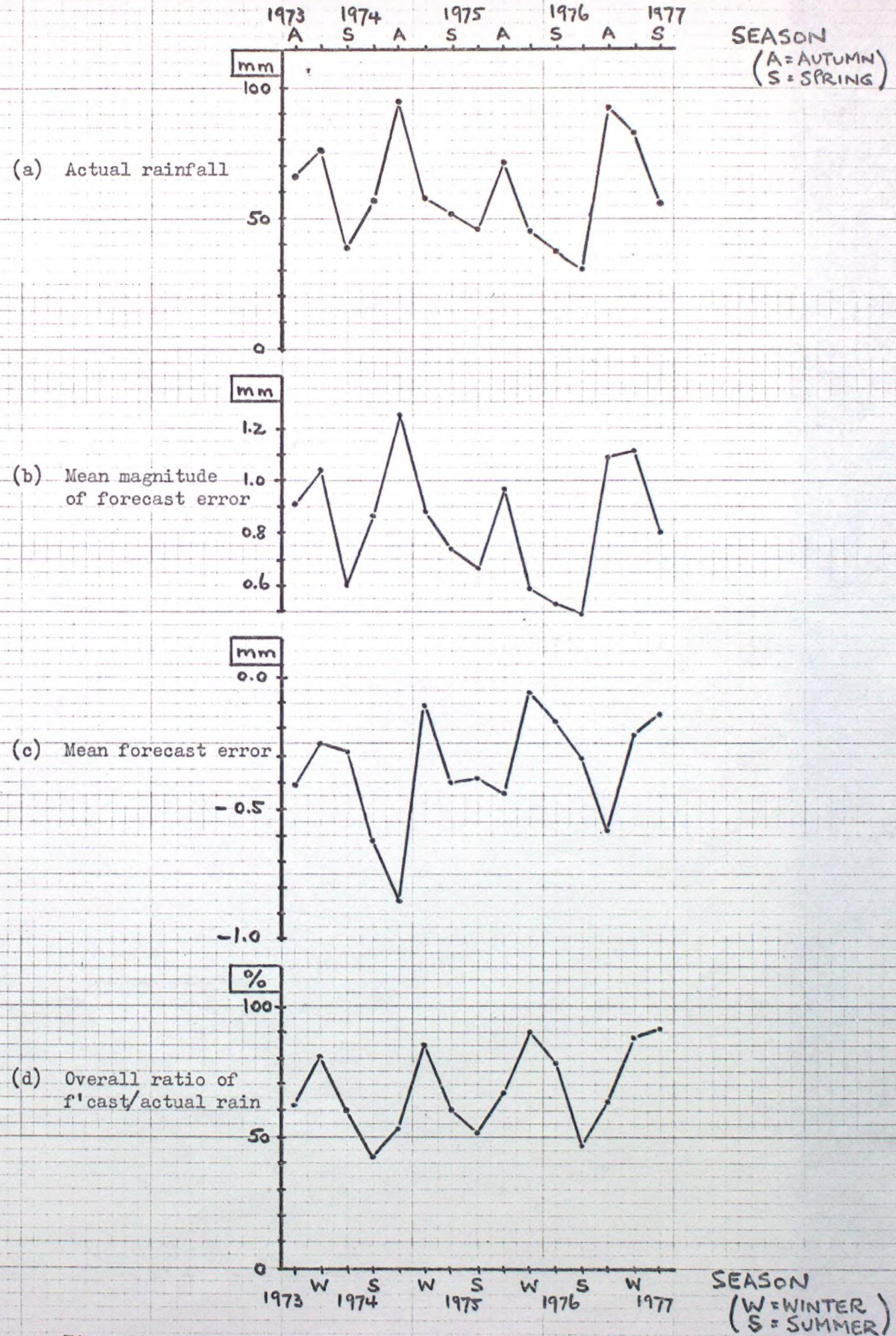

Figure 4.

Seasonal values of areal mean statistics for the whole region.

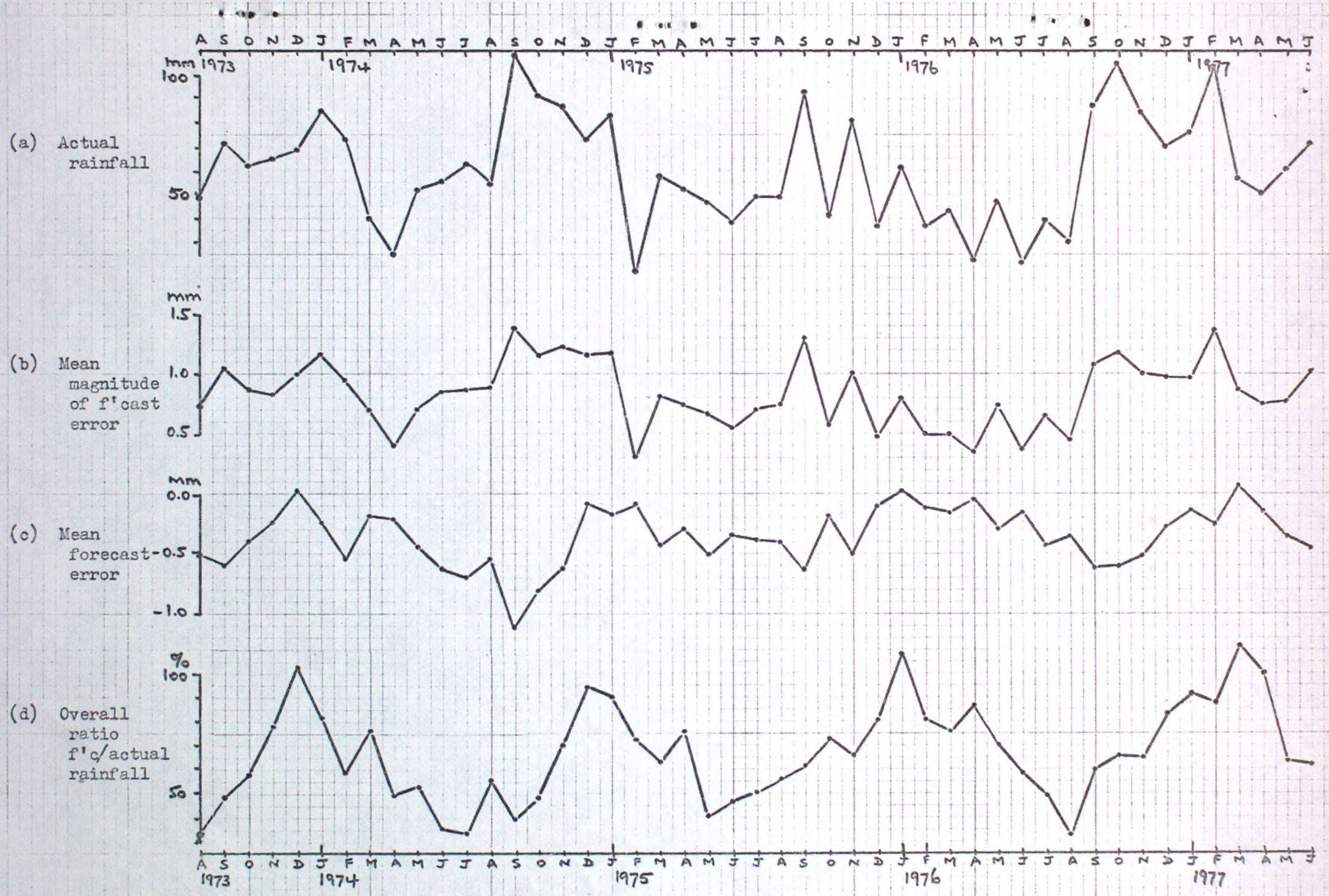

Figure 5. Monthly values of areal mean statistics for the whole region.

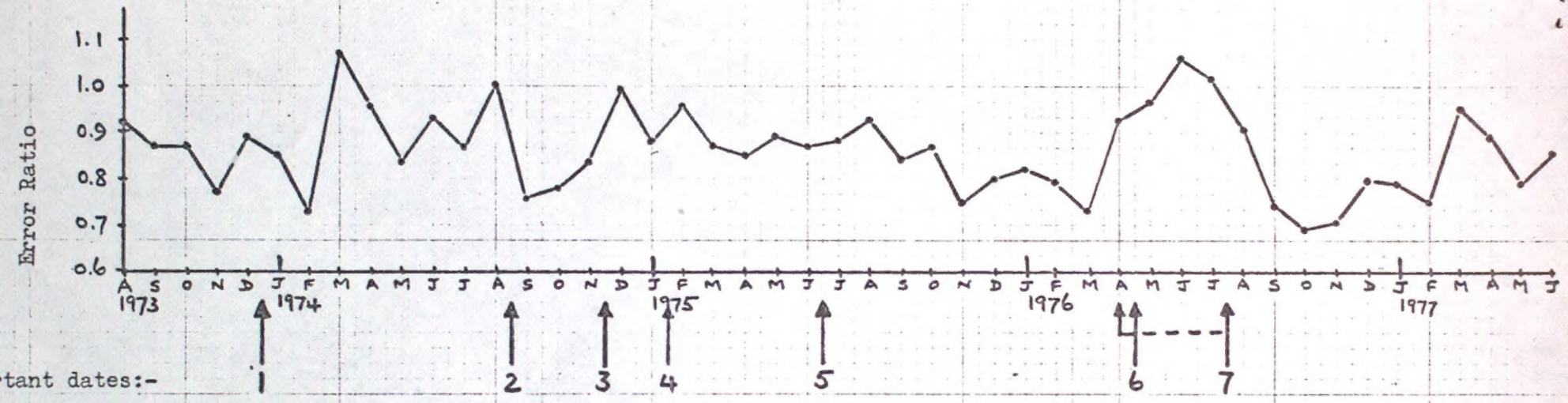

Figure 6. Monthly values of the Error Ratio for the whole region, and dates of some important events.

The Error Ratio is the total magnitude of the aggregated monthly errors, expressed as a proportion of the actual monthly rainfall.

Key to dates:-

1. Ocean Weather Ship 'C' withdrawn from operation.
2. Rectangle forecasts are run before the octagon forecasts.
3. Intervention staff in C F O become operational.
4. Rectangle background fields interpolated from octagon.
5. (a) Ocean Weather Ship 'C' reinstated.
   (b) Deep convection scheme introduced into rectangle forecasts.
6. New lateral boundary scheme for the rectangle.
7. Incorrect humidity background fields in rectangle analyses.